Full wwPDB X-ray Structure Validation Report (i)

Dec 21, 2021 – 04:08 PM JST

PDB ID	:	7CQ2
Title	:	Crystal structure of Slx1-Slx4
Authors	:	Xu, X.; Wang, M.; Sun, J.; Li, G.; Yang, N.; Xu, R.M.
Deposited on	:	2020-08-08
Resolution	:	2.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

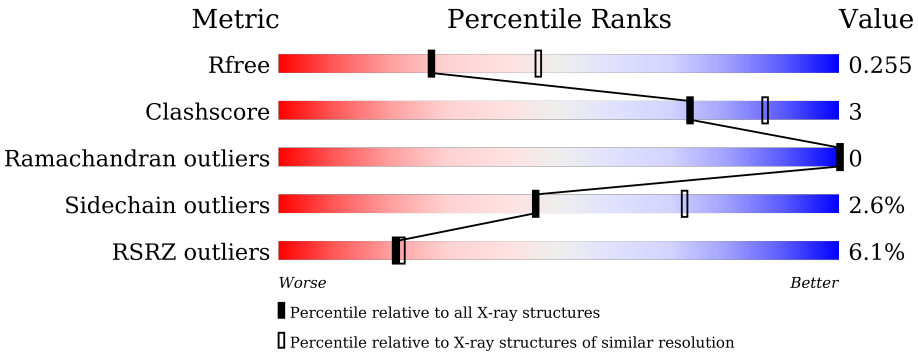
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.25
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.25

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	304	5% 81% 9% • 10%
1	B	304	8% 86% 12% •
2	C	150	% 48% • 51%
2	D	150	2% 43% • • 53%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6086 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Structure-specific endonuclease subunit SLX1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	0	1	0
			2283	1460	407	398	18			
1	B	297	Total	C	N	O	S	0	1	0
			2463	1573	445	427	18			

- Molecule 2 is a protein called SLX4 isoform 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	74	Total	C	N	O	S	0	1	0
			613	399	94	118	2			
2	D	71	Total	C	N	O	S	0	0	0
			586	383	90	111	2			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	598	MET	-	initiating methionine	UNP A0A6A5PU22
C	599	GLY	-	expression tag	UNP A0A6A5PU22
C	600	SER	-	expression tag	UNP A0A6A5PU22
C	601	SER	-	expression tag	UNP A0A6A5PU22
C	602	HIS	-	expression tag	UNP A0A6A5PU22
C	603	HIS	-	expression tag	UNP A0A6A5PU22
C	604	HIS	-	expression tag	UNP A0A6A5PU22
C	605	HIS	-	expression tag	UNP A0A6A5PU22
C	606	HIS	-	expression tag	UNP A0A6A5PU22
C	607	HIS	-	expression tag	UNP A0A6A5PU22
C	608	SER	-	expression tag	UNP A0A6A5PU22
C	609	GLN	-	expression tag	UNP A0A6A5PU22
D	598	MET	-	initiating methionine	UNP A0A6A5PU22
D	599	GLY	-	expression tag	UNP A0A6A5PU22
D	600	SER	-	expression tag	UNP A0A6A5PU22
D	601	SER	-	expression tag	UNP A0A6A5PU22

Continued on next page...

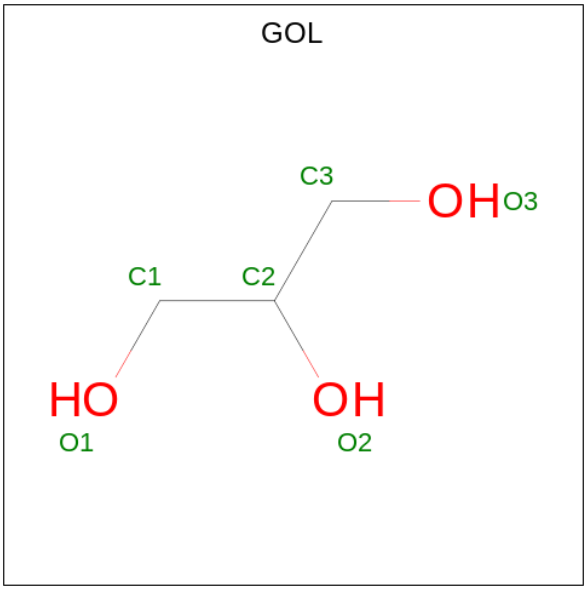
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	602	HIS	-	expression tag	UNP A0A6A5PU22
D	603	HIS	-	expression tag	UNP A0A6A5PU22
D	604	HIS	-	expression tag	UNP A0A6A5PU22
D	605	HIS	-	expression tag	UNP A0A6A5PU22
D	606	HIS	-	expression tag	UNP A0A6A5PU22
D	607	HIS	-	expression tag	UNP A0A6A5PU22
D	608	SER	-	expression tag	UNP A0A6A5PU22
D	609	GLN	-	expression tag	UNP A0A6A5PU22

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total Zn 2 2	0	0
3	B	2	Total Zn 2 2	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0

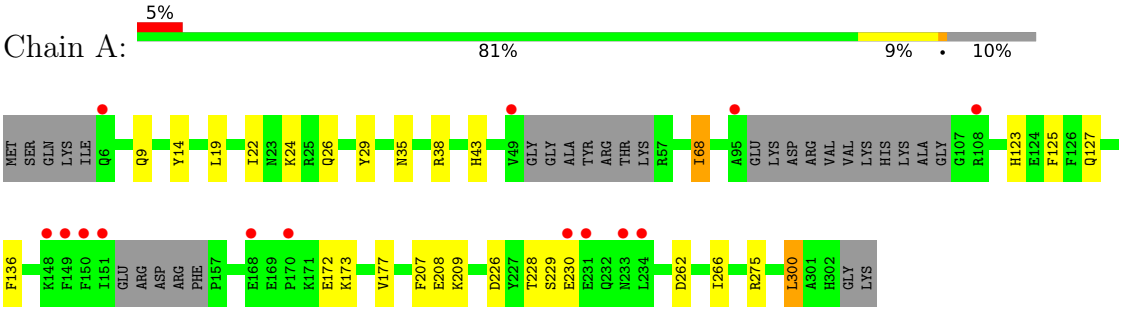
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	58	Total O 58 58	0	0
5	B	36	Total O 36 36	0	0
5	C	21	Total O 21 21	0	0
5	D	10	Total O 10 10	0	0

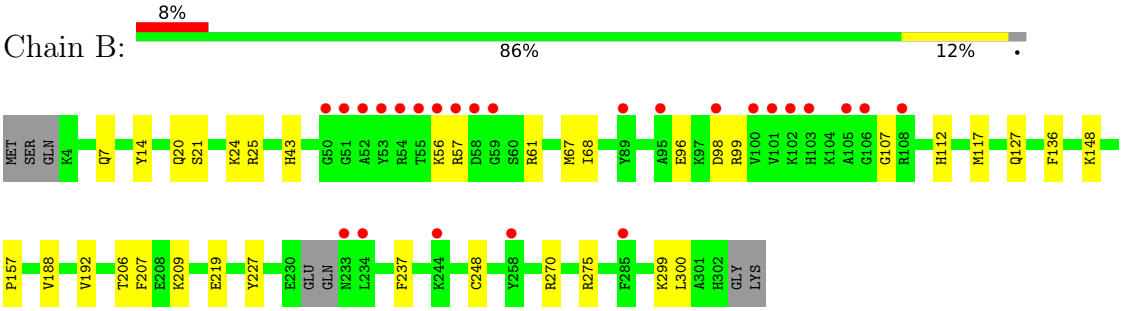
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

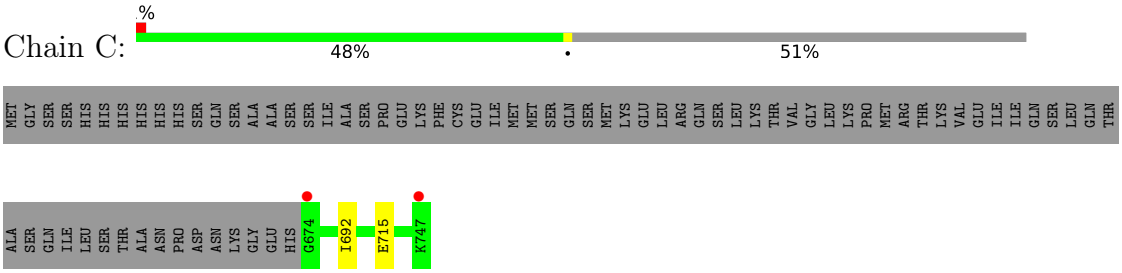
• Molecule 1: Structure-specific endonuclease subunit SLX1



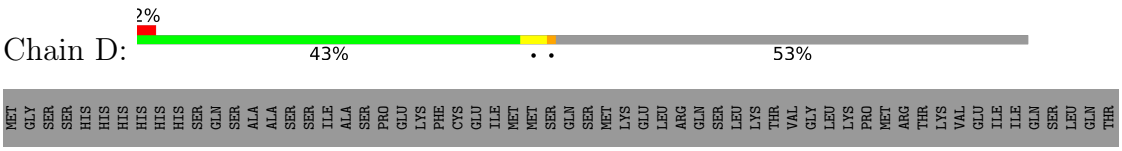
• Molecule 1: Structure-specific endonuclease subunit SLX1

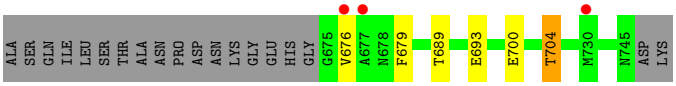


• Molecule 2: SLX4 isoform 1



• Molecule 2: SLX4 isoform 1




ALA
SER
GLN
ILE
LEU
SER
THR
ALA
ASN
PRO
ASP
ASN
LYS
GLY
GLU
HIS
GLY
G675
V676
A677
N678
F679
T689
E693
E700
T704
M730
N745
ASP
LYS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.28Å 75.96Å 186.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.49 – 2.50 42.49 – 2.47	Depositor EDS
% Data completeness (in resolution range)	98.4 (42.49-2.50) 98.2 (42.49-2.47)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.07 (at 2.48Å)	Xtriage
Refinement program	PHENIX 1.7.3_928	Depositor
R, R_{free}	0.201 , 0.262 0.196 , 0.255	Depositor DCC
R_{free} test set	1616 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	45.7	Xtriage
Anisotropy	0.661	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 36.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6086	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.86% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, GOL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.28	0/2338	0.42	0/3146
1	B	0.26	0/2523	0.41	0/3394
2	C	0.27	0/630	0.44	0/854
2	D	0.28	0/600	0.45	0/814
All	All	0.27	0/6091	0.42	0/8208

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2283	0	2273	15	0
1	B	2463	0	2465	20	0
2	C	613	0	596	1	0
2	D	586	0	570	4	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
4	A	6	0	8	1	0
4	B	6	0	8	2	0
5	A	58	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	36	0	0	0	0
5	C	21	0	0	0	0
5	D	10	0	0	1	0
All	All	6086	0	5920	39	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (39) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:43:HIS:HA	4:B:403:GOL:H2	1.68	0.76
1:A:43:HIS:HA	4:A:403:GOL:H2	1.70	0.72
1:A:35:ASN:HB3	1:A:38:ARG:HH21	1.67	0.60
1:A:172:GLU:HB2	1:A:177:VAL:HG21	1.85	0.59
1:B:136:PHE:HE1	1:B:192:VAL:HG21	1.69	0.57
2:D:700:GLU:O	2:D:704:THR:HB	2.05	0.56
1:A:22:ILE:O	1:A:173:LYS:NZ	2.38	0.56
1:A:123:HIS:CD2	1:A:125:PHE:H	2.24	0.55
1:B:270:ARG:NH1	2:D:693:GLU:OE2	2.40	0.54
1:B:96:GLU:HA	1:B:99:ARG:HH11	1.71	0.53
1:B:127:GLN:OE1	1:B:157:PRO:HG2	2.10	0.51
1:B:107:GLY:HA2	1:B:112:HIS:CG	2.45	0.51
1:A:9:GLN:HB2	1:A:300:LEU:HD13	1.92	0.50
1:B:68:ILE:HG23	1:B:136:PHE:HE2	1.75	0.50
1:B:219:GLU:HG3	1:B:248:CYS:HB2	1.94	0.49
1:A:226:ASP:OD1	1:A:228:THR:OG1	2.26	0.49
1:B:7:GLN:HB2	1:B:300:LEU:HD12	1.94	0.49
1:A:35:ASN:HB3	1:A:38:ARG:NH2	2.26	0.48
1:A:68:ILE:HG23	1:A:136:PHE:HE2	1.80	0.47
1:B:56:LYS:HG3	1:B:57:ARG:H	1.81	0.46
1:B:206:THR:HA	1:B:209:LYS:HE2	1.98	0.46
1:A:19:LEU:HD11	1:A:43:HIS:CD2	2.51	0.46
1:B:21:SER:HB3	1:B:24:LYS:O	2.16	0.45
1:A:24:LYS:HB3	1:A:26:GLN:NE2	2.32	0.44
1:A:19:LEU:HB2	1:A:29:TYR:HB3	1.99	0.43
2:C:692:ILE:HD13	2:C:692:ILE:HA	1.88	0.43
1:B:25:ARG:HD3	1:B:25:ARG:HA	1.88	0.43
2:D:704:THR:HG21	5:D:801:HOH:O	2.17	0.42
1:B:237:PHE:CE1	1:B:299:LYS:HD3	2.53	0.42
1:A:229:SER:OG	1:A:230:GLU:N	2.50	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:61:ARG:NH2	4:B:403:GOL:O2	2.48	0.42
1:B:107:GLY:HA2	1:B:112:HIS:CD2	2.54	0.42
1:B:136:PHE:CD1	1:B:188:VAL:HG12	2.55	0.42
2:D:689:THR:O	2:D:693:GLU:HG3	2.21	0.41
1:A:127:GLN:HE21	1:A:127:GLN:HB2	1.66	0.40
1:B:117:MET:HE3	1:B:117:MET:HB3	1.72	0.40
1:A:262:ASP:O	1:A:266:ILE:HG12	2.22	0.40
1:B:20:GLN:NE2	1:B:25:ARG:HD2	2.37	0.40
1:B:67:MET:SD	1:B:117:MET:HE1	2.61	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	267/304 (88%)	261 (98%)	6 (2%)	0	100	100
1	B	294/304 (97%)	285 (97%)	9 (3%)	0	100	100
2	C	73/150 (49%)	73 (100%)	0	0	100	100
2	D	69/150 (46%)	69 (100%)	0	0	100	100
All	All	703/908 (77%)	688 (98%)	15 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	253/276 (92%)	245 (97%)	8 (3%)	39	65
1	B	270/276 (98%)	264 (98%)	6 (2%)	52	77
2	C	68/135 (50%)	67 (98%)	1 (2%)	65	85
2	D	65/135 (48%)	62 (95%)	3 (5%)	27	50
All	All	656/822 (80%)	638 (97%)	18 (3%)	46	71

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	TYR
1	A	68	ILE
1	A	207	PHE
1	A	208	GLU
1	A	209	LYS
1	A	275[A]	ARG
1	A	275[B]	ARG
1	A	300	LEU
1	B	14	TYR
1	B	98	ASP
1	B	148	LYS
1	B	207	PHE
1	B	227	TYR
1	B	275	ARG
2	C	715	GLU
2	D	676	VAL
2	D	679	PHE
2	D	704	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	123	HIS
1	A	127	GLN
1	A	161	GLN
1	A	242	ASN
1	B	90	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	GOL	B	403	-	5,5,5	0.35	0	5,5,5	0.20	0
4	GOL	A	403	-	5,5,5	0.37	0	5,5,5	0.18	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	B	403	-	-	2/4/4/4	-
4	GOL	A	403	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	403	GOL	O1-C1-C2-C3
4	B	403	GOL	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	A	403	GOL	O1-C1-C2-O2
4	B	403	GOL	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	403	GOL	2	0
4	A	403	GOL	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	274/304 (90%)	0.11	14 (5%)	28	29	34, 48, 74, 89	0
1	B	297/304 (97%)	0.45	25 (8%)	11	11	38, 59, 87, 98	0
2	C	74/150 (49%)	-0.03	2 (2%)	54	58	35, 45, 62, 73	0
2	D	71/150 (47%)	-0.05	3 (4%)	36	39	42, 49, 59, 70	0
All	All	716/908 (78%)	0.22	44 (6%)	21	22	34, 51, 81, 98	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	150	PHE	7.2
1	B	100	VAL	5.6
1	A	149	PHE	5.5
1	B	234	LEU	5.0
1	B	53	TYR	4.6
1	B	101	VAL	4.5
1	B	58	ASP	4.1
1	B	108	ARG	4.0
1	A	6	GLN	3.9
1	A	95	ALA	3.8
1	B	95	ALA	3.8
2	C	674	GLY	3.8
1	B	105	ALA	3.6
1	B	106	GLY	3.6
1	B	57	ARG	3.5
1	B	52	ALA	3.5
2	D	676	VAL	3.5
1	B	56	LYS	3.4
1	B	51	GLY	3.4
2	C	747	LYS	3.3
1	B	50	GLY	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	148	LYS	3.0
1	B	54	ARG	3.0
1	B	103	HIS	2.9
1	A	230	GLU	2.8
1	B	55	THR	2.8
1	B	102	LYS	2.7
1	B	244	LYS	2.7
1	B	89	TYR	2.6
1	B	98	ASP	2.6
2	D	677	ALA	2.5
1	A	151	ILE	2.5
2	D	730	MET	2.5
1	B	233	ASN	2.5
1	A	231	GLU	2.4
1	B	285	PHE	2.4
1	A	170	PRO	2.4
1	B	59	GLY	2.4
1	A	108	ARG	2.3
1	A	49	VAL	2.1
1	B	258	TYR	2.0
1	A	168	GLU	2.0
1	A	233	ASN	2.0
1	A	234	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	GOL	B	403	6/6	0.81	0.20	59,66,70,71	0
4	GOL	A	403	6/6	0.88	0.17	49,51,51,51	0
3	ZN	B	402	1/1	0.98	0.09	72,72,72,72	0
3	ZN	B	401	1/1	0.99	0.08	71,71,71,71	0
3	ZN	A	401	1/1	0.99	0.14	45,45,45,45	0
3	ZN	A	402	1/1	1.00	0.10	46,46,46,46	0

6.5 Other polymers (i)

There are no such residues in this entry.